

CLAIMS

What is claimed is:

1. A method of supporting a transaction application and a parallel application across a plurality of remote clustered systems based on a service level agreement, comprising:

monitoring a performance of a local clustered system in response to the transaction application, based on the service level agreement and a workload of the local clustered system;

analyzing the performance of the local clustered system to identify a violation of the service level agreement, if any, by the local clustered system;

in response to the identified violation, dynamically reallocating a local computing resource, if available, from the parallel application to the transaction application that requires an additional computing resource to meet the service level agreement;

if the local computing resource is not available, sending a request for allocation of a remote computing resource to at least one of the remote clustered systems.

2. The method of claim 1, wherein the parallel application comprises a numerically intensive application.

3. The method of claim 1, wherein the transaction application comprises a plurality of discrete events that are less numerically intensive than the parallel application.

4. The method of claim 1, wherein the local clustered system comprises a cluster of computers that process the transaction application and the parallel application.
5. The method of claim 1, wherein the remote clustered systems comprise clusters of computers that process transaction applications and parallel applications.
6. The method of claim 1, wherein the remote clustered systems are geographically co-located with the local clustered system.
7. The method of claim 1, wherein the remote clustered systems are geographically remote from the local clustered system.
8. The method of claim 1, wherein the service level agreement defines an acceptable performance of the local clustered system in response to the transaction application.
9. The method of claim 1, wherein the service level agreement defines an acceptable performance of the local clustered system in response to the parallel application.
10. The method of claim 1, further comprising making a prediction of the performance of the local clustered system to identify a potential violation of the service level agreement, if any, by the performance of the local clustered system.
11. The method of claim 1, wherein the violation comprises an actual violation of the service level agreement by the performance of the local clustered system

12. The method of claim 10, wherein the violation comprises a predicted violation of the service level agreement by the performance of the local clustered system.

13. The method of claim 12, wherein the remote computing resource comprises an under-utilized computing resource.

14. The method of claim 12, wherein the remote computing resource comprises a computing resource that is processing a lower-priority workload.

15. The method of claim 1, wherein sending the request for allocation of the remote computing resource comprises sending a value proposed to be paid by the local clustered system in compensation for the remote computing resource.

16. The method of claim 15, wherein the value is based on a violation cost to the local clustered system for the violation of the service level agreement.

17. The method of claim 1, further comprising the remote clustered system determining whether the request can be fulfilled based on a service level agreement and a workload of the remote clustered system, and a cost utility.

18. The method of claim 17, wherein the cost utility comprises an allocation cost to the remote clustered system for allocating the remote computing resource to the local clustered system based on a possibility of violating the service level agreement of the remote clustered system.

19. The method of claim 18, wherein if the value is less than the cost utility, the remote clustered system sending a counter offer based on the cost utility to the local clustered system.

20. The method of claim 18, wherein if the value is less than the cost utility, the remote clustered system sending a rejection of the request to the local clustered system.

21. The method of claim 18, further comprising the remote clustered system agreeing to fulfill the request for the remote computing resource.

22. The method of claim 18, further comprising a plurality of remote clustered systems agreeing to fulfill the request for the remote computing resource.

23. The method of claim 22, further comprising the local clustered system identifies a selecting remote computing resource from the plurality of remote clustered systems to fulfill the request.

24. The method of claim 23, further comprising the local clustered system sending an acceptance to the selected remote clustered system to fulfill the request.

25. The method of claim 24, further comprising provisioning, if necessary, the remote computing resource to execute the transaction application.

26. The method of claim 24, further comprising provisioning, if necessary, the remote computing resource to execute the parallel application.

27. A computer program product having instruction codes for supporting a transaction application and a parallel application across a plurality of remote clustered systems based on a service level agreement, comprising:

- a first set of instruction codes for monitoring a performance of a local clustered system in response to the transaction application, based on the service level agreement and a workload of the local clustered system;

- a second set of instruction codes for analyzing the performance of the local clustered system to identify a violation of the service level agreement, if any, by the local clustered system;

- a third set of instruction codes, which, in response to the identified violation, dynamically reallocates a local computing resource, if available, from the parallel application to the transaction application that requires an additional computing resource to meet the service level agreement; and

- if the local computing resource is not available, the third set of instruction codes sending a request for allocation of a remote computing resource to at least one of the remote clustered systems.

28. The method of claim 27, wherein the third set of instruction codes further sends, to the at least one of the remote clustered systems, a value proposed to be paid by the local clustered system in compensation for the remote computing resource.

29. A system for supporting a transaction application and a parallel application across a plurality of remote clustered systems based on a service level agreement, comprising:

a server allocation controller monitors a performance of a local clustered system in response to the transaction application, based on the service level agreement and a workload of the local clustered system;

a service level agreement monitor analyzes the performance of the local clustered system to identify a violation of the service level agreement, if any, by the local clustered system;

a server allocation manager which, in response to the identified violation, dynamically reallocates a local computing resource, if available, from the parallel application to the transaction application that requires an additional computing resource to meet the service level agreement; and

if the local computing resource is not available, the local clustered system sends a request for allocation of a remote computing resource to at least one of the remote clustered systems.

30. The system of claim 29, wherein the local clustered system further sends, to the at least one of the remote clustered systems, a value proposed to be paid by the local clustered system in compensation for the remote computing resource.

31. A method for supporting a transaction application and a parallel application by a local clustered system that implements a service level agreement across, the method comprising:

specifying a performance parameter for the service level agreement;

invoking a server allocation utility, wherein the performance parameter is made available to the server allocation utility for reallocating computing resources on a local clustered system and on a plurality of remote systems, to meet the service level agreement;

wherein if a local computing resource is not available to the local clustered system to meet the service level agreement, requesting a remote computing resource from at least one of plurality of remote clustered systems to meet the service level agreement; and

receiving a level of performance by the local clustered system within the parameter of the service level agreement for a contracted execution of the transaction application and the parallel application, wherein in response to a violation of the service level agreement, the local clustered system server allocation utility dynamically reallocates the remote computing resource to any one of the transaction application and the parallel application that requires an additional computing resource.

32. The method of Claim 31, wherein the violation comprises an actual violation of the service level agreement by the performance of the clustered system.

33. The method of Claim 32, wherein the violation comprises a predicted violation of the service level agreement by the performance of the clustered system.